

修士学位論文

論文題目 ユーザの体験情報を用いた観光支援シス

テム

氏 名

はん
潘

けんた
健太

専 攻

情報学専攻

指導教授

北山 大輔 准教授

修了年月(西暦)

2020年3月

工学院大学大学院

要旨

目次

第1章	序論	1
第2章	本研究のアプローチ	5
2.1	観光支援システム	5
2.2	関連研究	6
第3章	ユーザの既訪問スポットの位置付けに基づく未訪問スポットの説明手法	9
3.1	概要	9
3.2	未訪問スポットの説明手法	9
3.3	キーワード抽出に関する評価	9
3.4	対応付けに関する評価	9
3.5	未訪問スポット説明の有効性評価	9
第4章	地図上における未訪問スポットの説明性向上のための観光スポットの対応関係可視化手法	10
4.1	概要	10
4.2	観光スポット間の対応関係を用いた可視化	10
4.3	対応関係の可視化評価	10
第5章	考察	11
第6章	結論	12
	謝辞	13
	本研究に関する発表	14
	その他の発表実績	15
	受賞	16

図目次

1.1	レビューサイト上で観光スポットに付与される情報例	2
2.1	観光支援システムの流れ	6

表目次

第 1 章

序論

一般に，ユーザは観光旅行する前に Web 上の情報やガイドブックを活用して計画を立てることが多くなっている．その際，目的とする観光スポットに関する概要やアクセスするための方法などの情報を集める．ガイドブックは，有名な観光スポットや流行りのものに焦点を当て，そのスポットの料金や見どころ，そして食べ物などが掲載されており，その土地に初めて観光で訪れるユーザや，普段から頻繁に観光旅行をしないユーザにとって満足を得られるような内容が書かれている．そのため，ガイドブックは普段から頻繁に観光旅行をするようなユーザや，その観光スポットに対してこだわりの観点を持つユーザの個別の嗜好までは考慮できない．

Web 上の情報に関して，観光スポットに関する総合的な情報が掲載された公式サイト，過去にそこを訪れた観光者による旅行記ブログサイト，そして，評価が記載されたレビューサイトなどの様々なサイトがある．しかしながら，種々のサイトが混在しており，膨大な情報から自分の嗜好に合う観光スポットの情報を得ることは簡単ではない．目的の情報を収集するためには手動で閲覧し取捨選択しなければならない．旅行計画を立つために，まず訪問したいエリアを決定する必要がある．その後，観光スポット検索サイトやガイドブックの観光スポット情報を調べて観光スポットを決定する．ユーザの意識決定の材料の 1 つとして，るぶトラベル^{*1}，Tripadvisor^{*2}やじゃらん^{*3}などの観光スポット検索サイトがある．これらのサイトには特定の観光スポットを訪問したことのあるユーザがレビューを投稿し，観光スポットに関する豊富な情報が存在している．観光スポット検索サイト上で観光スポットに付与される情報の例^{*4}を図??に示す．なお，これらの情報のうちユーザレビュー以外の情報(概要文・タグ・評価点など)はメタデータと呼ばれる．

観光スポット検索サイトに掲載された情報を基に，利用者(ユーザ)は自身の嗜好や目的

^{*1} <https://www.rurubu.travel/>

^{*2} <https://www.tripadvisor.com/>

^{*3} <https://www.jalan.net/kankou/>

^{*4} <https://www.jalan.net/kankou/130000/?screenId=OUW2201>

図 1.1 レビューサイト上で観光スポットに付与される情報例

に合致した特徴を持つ対象観光スポットを検索する。多くの観光スポット検索サイトでは、対象観光スポットに付与されるメタデータを利用した以下のような検索方法が提供されている。

- ユーザが指定したタグを持つ対象観光スポット
- 評価点などの数値メタデータに基づく並び替え
- 入力単語(クエリ)を概要文中に含むオブジェクトの検索

しかし、これらのメタデータを用いた検索方法だけでは検索が困難な場合も存在する。たとえば、対象観光スポットを実際に利用することでユーザが得られる印象・感情・体験などを表す特徴に基づき対象観光スポットを検索したい場合などが該当する。また、訪問したいエリア内に数多くの観光スポットが存在している。ユーザは検索エリアに関する事前知識が

ないため、検索された多数の観光スポットからどの観光スポットのレビューを読むべきか効率的に判断することは困難で、自身のイメージから外れない観光スポットを見つけることは容易ではない。そこで、本研究では、さまざまな観光スポットを効果的に理解するためには、ユーザが訪問した体験のある観光スポットとを使って未訪問スポットを検索する手法を提案した。観光支援システムを構築するにあたり課題となる点は2つ説明する。

1. 観光スポット間の対応付け

ある観光スポットに対応する観光スポットを判定する手法として、レビューやカテゴリが類似したコンテンツであるという観点で判定する方法が一般的である。観光スポットで用いられる単語に対する従来の特徴量 TFIDF は、観光スポット内での重要な単語を判定する際に用いられる。しかし、従来手法の TFIDF はユーザの観光スポット選択における意図や他スポットへの関心について考慮されていない。そのため、ある観光スポットが他の観光スポットと比較するとき、その観光スポットの特徴を表すことができない。そこで、あるスポットに対し、既に訪問した観光スポットと比較した相対的特徴ベクトルを作成し、相対的特徴ベクトル間の類似度に基づいて既訪問スポットと未訪問スポット間の対応付け手法を提案する。第3章で、より具体的な提案手法について述べる。

2. 観光スポットの対応関係を地図上に可視化

地理的情報の可視化を意味的な関係の可視化に取り込む情報の可視化としては、抽出した情報を地図上にマッピングするのが一般的である。また、意味的な関係性の可視化として、グラフモデル上でオブジェクト間のつながりを関連度で表現することが多い。観光スポット間の意味的な対応関係を可視化するために、未訪問スポットは地図上で実在する座標に固定され、配置の自由があるのは既訪問スポットのみという制約がある。そこで、ある未訪問スポットに対して、入力した既訪問スポットがどのような関係にあるのか地図上で、一目で把握できるために、既訪問スポットと未訪問スポットの対応関係を可視化する手法を提案する。第4章で、より具体的な提案手法について述べる。

以下に本論文の構成を示す。まず、2章で本研究のアプローチについて述べる。ここでは観光支援システムのアルゴリズムと関連研究について述べる。この観光支援システムのための2つの観点に基づく各手法は3章と4章で述べる。3章では、ユーザの既訪問スポットの位置付けに基づく未訪問スポットの説明手法について述べる。4章では、地図上における未訪問スポットの説明向上のための観光スポットの対応付け可視化手法について述べる。5章

では、各手法に対する評価実験結果を踏まえた提案手法全体に対する考察を行う。6章では、まとめと今後の課題を述べる。

第 2 章

本研究のアプローチ

2.1 観光支援システム

観光支援システムの 1 つに、求める観光スポット情報を発見するために観光スポットの理解をしやすくすることが考えられる。観光情報収集までの過程を効率化することで、ユーザの観光計画への敷居は下がり、訪れたことない観光スポットの理解の向上も望める。本節では、ユーザいままで体験したことがありかつ気に入った観光スポットを利用し、訪問したいエリア内の観光スポットに対して関連付けすることで、手軽に観光情報を理解する観光支援システムを提案する。観光支援システムの具体的な流れを図 2.1 に沿って説明する。

まず、ユーザの体験履歴の中ていままで訪れたことがあるかつ気に入った観光スポットを入力する。次に、これから行ってみたい都道府県やエリアを入力する。そして、データベースで保存されている 2016 年 10 月までのじゃらんの観光スポットレビューデータを使って提案手法に適用することで、ユーザの未知な観光スポットに対する理解を支援するため、既に訪問したことがある観光スポットの特徴を用いて、未訪問エリアの観光スポットを説明するために対応付けを行う。

本研究における観光スポットの対応付けに関して、ユーザによって、その強調すべき特徴が変化する。たとえば、日本の有名な観光スポットを履歴に持つユーザにとっての「東京都庁舎展望台」と「金閣寺」について考える。このとき、「金閣寺」の相対的特徴は、お寺、金色、京都などである。「東京都庁舎展望台」の相対的特徴は、展望、夜景、建物、新宿などが考えられる。

一方、京都の寺院を履歴に持つユーザにとっての「金閣寺」と「清水寺」について考える。このとき、同じ「金閣寺」でもお寺や京都などは特徴にはならず、「金閣寺」の相対的特徴は、金色、金箔、輝きなどとなる。「清水寺」の相対的特徴は、舞台や一望などである。

対応付けを行ったなあと、ユーザに未訪問スポットを説明するための観光スポット特徴語を抽出する。最後に、地図上で既訪問スポットと未訪問スポットの対応関係と特徴語を可視化することで、ユーザが未訪問スポットに対する理解を容易化することを目指す。

図 2.1 観光支援システムの流れ

2.2 関連研究

2.2.1 体験履歴による検索や推薦システムに関する研究

倉島ら [?] は, Flickr に投稿された写真のジオタグ情報を人々の旅行履歴として利用した旅行ルート推薦手法を提案した. この手法では, ユーザの現在地から行きやすい場所とユーザの興味に合致した場所に移動しやすいと仮定し, 行動モデルを生成している. ユーザのジオタグ付き写真集合は, 時間情報でソートすると個人の旅行履歴とみなすことができると考え, ジオタグ情報を利用してユーザの行動モデルを生成している.

Cheng ら [?] は, 自由に利用できるコミュニティ投稿の写真を活用して, パーソナライズされた旅行のおすすめに焦点を当て, 特定のユーザプロフィールまたは属性を考慮し, パー

ソナライズされた旅行の推奨を行うことを提案した。

観光地検索するとき、松本ら [?] はクチコミから特徴語を抽出して利用する研究を行った。抽出対象を任意の名詞として、4 種類の手法、TFIDF, ATF(Average Term Frequency), ポアソン確率, エントロピーのうちどの手法が特徴語抽出に適しているのか検討を行った。また、抽出した特徴語を利用した検索支援システムを試作し、実験を通して特徴語提示の効果を検証した。

嶋田ら [?] は Web から観光情報を抽出し、複数の特徴ベクトルから観光地間の類似性を評価することで、観光地を推薦するシステムを提案した。観光地の特徴ベクトルは、知恵袋・ブログ上での共起キーワードと時系列分布、知恵袋上でのカテゴリ構造、観光地周辺施設、地図画像から生成した。これらの特徴ベクトルから観光地間の類似度の測定を行い、類似度の高い観光地を推薦した。

廣嶋ら [?] は地理情報検索の際のクエリ入力支援として、提示する特徴語の抽出手法について研究を行った。この手法では、各ブログ記事から特徴語候補の抽出および地点の特定を行った。具体的には、特徴語の候補を Wikipedia の見出し語に限定し、ポアソン確率を用いて特徴語抽出を行った。また、地点の特定のための地名表現の抽出および緯度経度への変換の方法としては平野ら [?] の手法を用いた。

野守ら [?] は日本全国の観光地のクチコミデータを用いて、観光客が話題にする観光テーマを確率的に抽出し、そのテーマを軸として各観光地の特徴を定量的に評価した。また、クチコミのテキストデータにテキストマイニングを実行して表現を抽出し、観光地ごとにその表現の出現頻度を集計したクロス集計表に PLSA を実行することで、観光客のクチコミだけに基づいた観光テーマの抽出と観光地の特徴分析を行なった。

このように多くの研究はユーザに合致する観光スポットの抽出や、一般的な特徴の抽出に焦点をあてているが、本研究では、ユーザの選択性を高めるために、個人化した特徴を抽出する点では異なる。

佐藤 [?] はコスメアイテムに対してのクチコミからアイテムの特徴を抽出した。また、それぞれのユーザのクチコミからアイテムに対する不満に相当する項目を抽出し、これらを基に各アイテムの特徴を可視化することで不満を解消するような新しい商品を提案した。

林ら [?] は、レビュー中の評価表現とその評価対象に基づき、ユーザが書いた映画レビューから他者が書いた映画レビューを推薦する手法を提案した。評価表現とその評価対象に基づく推薦を実現するために、レビュー中の「良い」や「良くない」などの肯定的または否定的な評価表現とその評価対象に着目した。

服部ら [?] はレビューを分析して、その価値観と繋がり深い要素としてユーザの「こだ

わり」に着目し、その結果に基づくユーザモデリング手法について提案した。

従来のレビューを利用する手法では、クチコミを分析して、それぞれの嗜好や価値観を表す単語を重要視して、ユーザには単語レベルで提示されることが多い。単語レベルの提示では、レビューの文章の雰囲気損なうため、ユーザの入力としては、体験したことかつ気に入った観光スポットの全レビューを利用すると考えた。

2.2.2 情報可視化に関する研究

本研究では、地理的な情報の可視化と意味的な関係の可視化に取り込む。地理的な情報の可視化としては、抽出した情報を地図上にマッピングするのが一般的であり、本研究もそれに従っている。代表的な研究を以下に紹介する。櫻川ら [?] は、ソーシャルメディア上にアップロードされた写真のジオタグ情報と撮影時刻に基づいて写真の撮影者を分類する手法を提案した。分類された撮影者(観光者, 在住者)ごとにホットスポットを可視化した。また、ジオタグ情報と撮影時刻以外にテキストタグを加えて、写真が撮影された地域で行われるイベントの穴場スポットを発見し、地図上に提示する手法を提案した [?]. 倉田ら [?] は、新しい観光情報サービスの形として「他の旅行者がどこで撮影したいか」というデータから観光ポテンシャルを可視化し、地図上に表示する方法を提案した。

意味的な関係性の可視化としては、グラフモデル上でオブジェクト間のつながりを関連度を表現することが多い。代表的な研究を以下に紹介する。Kitamura ら [?] は、一般的な物体認識を用いて、過去の個人旅行写真から推定したユーザの旅行の嗜好に基づき観光地を推薦する方法を提案した。物体認識システムを用いて、写真で撮った被写体情報のキーワードを取得し、グラフ視覚化技術によってキーワードの共起を表現した。また、グラフの視覚化技術に基づいて旅行写真付きのグラフを視覚化するユーザインターフェイスを紹介した。上村ら [?] は、ユーザが投稿したタグ付き画像を用いたファッションスタイルの関係性の可視化手法を提案した。ファッションスタイルは類似するタグを空間座標に固定することによって関係を表している。

本研究では、観光スポット間の意味的な対応関係を可視化するために、未訪問スポットは地図上で実在する座標に固定され、配置の自由があるのは既訪問スポットのみという制約がある。

第 3 章

ユーザの既訪問スポットの位置付けに基づく未訪問スポットの説明手法

- 3.1 概要
- 3.2 未訪問スポットの説明手法
- 3.3 キーワード抽出に関する評価
- 3.4 対応付けに関する評価
- 3.5 未訪問スポット説明の有効性評価

第 4 章

地図上における未訪問スポットの説明性向上のための観光スポットの対応関係可視化手法

4.1 概要

4.2 観光スポット間の対応関係を用いた可視化

4.3 対応関係の可視化評価

第 5 章

考察

第 6 章

結論

謝辞

修士課程での研究全般に渡ってご指導賜りました北山大輔准教授に厚く御礼申し上げます。

お忙しい中貴重な時間を割いて副指導教員および副査を引き受けてくださり，ご指導賜りました三木良雄教授に感謝致します。

同様に，お忙しい中貴重な時間を割いて副査を引き受けてくださり，ご指導賜りました〇〇教授に感謝致します。

研究室配属からの3年間に渡って，切磋琢磨し心の支えにもなってくれた歴代のインタラクティブメディア研究室メンバ全員に感謝致します。

最後に，名前を書き切ることはできませんが，両親をはじめ，私のこれまでの人生を支えてくださった全ての方に心から感謝いたします。ありがとうございました。

本研究に関する発表

論文

- [1] 潘健太, 北山大輔, “説明性向上のためのユーザレビューを用いた観光スポットの対応付け手法,” 情報処理学会論文誌データベース, 第 84 号, 2019.

査読付き会議

- [2] Kenta Han, Daisuke Kitayama, “An Explanation Method of Unfamiliar Tourist Spots based on Roles of User’s Familiar Spots,” Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2019, Hong Kong, pp358-362, 13-15 March, 2019.

研究会

- [3] 潘健太, 北山大輔, “ユーザの既訪問スポットの位置付けに基づく未訪問スポットの説明手法,” 第 11 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2019), H7-2, 2019
- [4] 潘健太, 北山大輔, “地図上における未訪問スポットの説明性向上のための観光スポットの対応関係可視化手法,” 観光情報学会第 20 回研究発表会, 2019

その他の発表実績

研究会

- [1] 潘健太, 北山大輔, “ユーザのレビュー選択に基づく観光スポット検索手法,” 第 10 回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2018), H1-2, 2018
- [2] 潘健太, 北山大輔, “観光スポット検索のためのユーザのレビュー選択と特徴抽出に関する考察,” 電子情報通信学会技術研究報告, vol. 118, no. 107, DE2018-5, pp. 21-24, 2018

受賞

査読付き会議

- [1] Kenta Han, Daisuke Kitayama, “An Explanation Method of Unfamiliar Tourist Spots based on Roles of User’s Familiar Spots,” Lecture Notes in Engineering and Computer Science: Proceedings of The International MultiConference of Engineers and Computer Scientists 2019, Hong Kong, pp358-362, 13-15 March, 2019.